

Name _____

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

- 1) One of the reasons that the Monitoring the Future (MTF) project was started was "to study changes in the beliefs, attitudes, and behavior of young people in the United States." Data are collected from 8th, 10th, and 12th graders each year. To get a representative nationwide sample, surveys are given to a randomly selected group of students. In Spring 2004, students were asked about alcohol, illegal drug, and cigarette use. Describe the W's, if the information is given. If the information is not given, state that it is not specified. 1) _____
- Who:
 - What:
 - When:
 - Where:
 - How:
 - Why:

In July 2013, the Federal Drug Administration approved a new version of a drug used to treat opium dependence. The old version of the drug had received complaints about a bitter taste, an aftertaste, and that it took a long time to dissolve. The goal of the new version was to get more patients to take the drug as prescribed by addressing these issues. In addition to these improvements, experimenters monitored the existence and types of side effects of the drug.

- 2) List the variables. Indicate whether each variable is categorical or quantitative. If the variable is quantitative, tell the units. 2) _____
- 3) Describe the W's, if the information is given: 3) _____
- Who:
 - What:
 - When:
 - Where:
 - How:
 - Why:

In June 2003 *Consumer Reports* published an article on some sport-utility vehicles they had tested recently. They reported some basic information about each of the vehicles and the results of some tests conducted by their staff. Among other things, the article told the brand of each vehicle, its price, and whether it had a standard or automatic transmission. They reported the vehicle's fuel economy, its acceleration (number of seconds to go from zero to 60 mph), and its braking distance to stop from 60 mph. The article also rated each vehicle's reliability as much better than average, better than average, average, worse, or much worse than average.

- 4) Describe the W's, if the information is given: 4) _____
- Who:
 - What:
 - When:
 - Where:
 - How:
 - Why:

5) Consider the following part of a data set:

5) _____

Age (years)	Sex	Only child?	Height (inches)	Weight (pounds)	Credit Hours	GPA	Major
21	Female	Yes	67.00	140.0	16	3.60	animal science
20	Female	No	62.00	130.0	18	3.86	biology
28	Female	No	64.00	188.0	21	3.25	psychology
21	Male	No	65.00	140.0	15	2.95	psychology
24	Female	No	67.00	130.0	20	3.00	anthropology
22	Male	Yes	68.00	135.0	15	2.94	journalism

List the variables in the data set. Indicate whether each variable is treated as categorical or quantitative in this data set. If the variable is quantitative, state the units.

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

6) What is the marginal distribution of breeds?

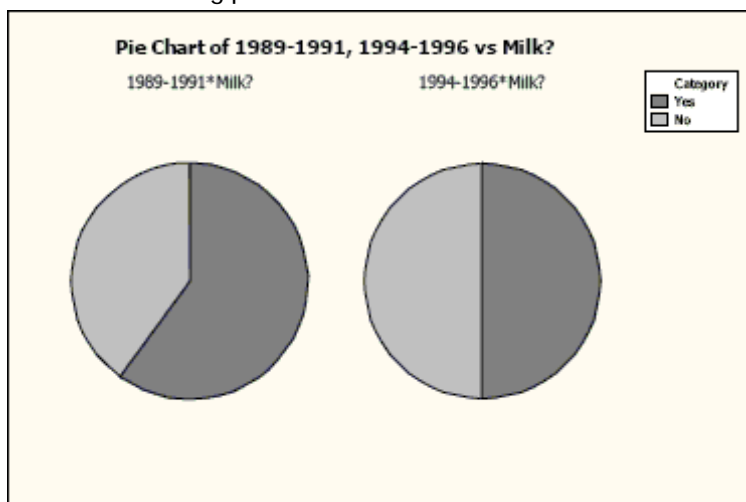
6) _____

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, 102(9), pp. 1234-1239):

		Nationwide Food Survey Years		
		1987-1988	1989-1991	1994-1996
Drinks Fluid Milk	Yes	354	502	366
	No	226	335	366
	Total	580	837	732

7) Consider the following pie charts of a subset of the data above:

7) _____



Do the pie charts above indicate that milk consumption by young girls is independent of the nationwide survey year? Explain.

In order to plan transportation and parking needs at a private high school, administrators asked students how they get to school. Some rode a school bus, some rode in with parents or friends, and others used "personal" transportation - bikes, skateboards, or just walked. The table summarizes the responses from boys and girls.

	Male	Female	Total
Bus	30	34	64
Ride	37	45	82
Personal	19	23	42
Total	86	102	188

8) What is the marginal distribution of gender?

8) _____

Has the percentage of young girls drinking milk changed over time? The following table is consistent with the results from "Beverage Choices of Young Females: Changes and Impact on Nutrient Intakes" (Shanthy A. Bowman, *Journal of the American Dietetic Association*, 102(9), pp. 1234-1239):

		Nationwide Food Survey Years			
		1987-1988	1989-1991	1994-1996	Total
Drinks Fluid Milk	Yes	354	502	366	1222
	No	226	335	366	927
	Total	580	837	732	2149

- 9) Do you think that milk consumption by young girls is independent of the nationwide survey year? Use statistics to justify your reasoning. 9) _____

To determine if people's preference in dogs had changed in the recent years, organizers of a local dog show asked people who attended the show to indicate which breed was their favorite. This information was compiled by dog breed and gender of the people who responded. The table summarizes the responses.

	Female	Male	Total
Yorkshire Terrier	73	59	132
Dachshund	49	47	96
Golden Retriever	58	33	91
Labrador	37	41	78
Dalmatian	45	28	73
Other breeds	86	67	153
Total	348	275	623

- 10) Do you think the breed selection is independent of gender? Give statistical evidence to support your conclusion. 10) _____

- 11) There is a proposal to replace the shortest roller coaster above with one that has a length of 1,200 ft. Indicate whether changing that roller coaster's length would make each of these summary statistics increase, decrease, or stay about the same. 11) _____
- mean
 - median
 - range
 - IQR
 - standard deviation

- 12) The students in a biology class kept a record of the height (in centimeters) of plants for a class experiment.

12) _____

49	67	38	55	62
54	36	41	56	43
48	75	44	60	48
52	48	53	59	32

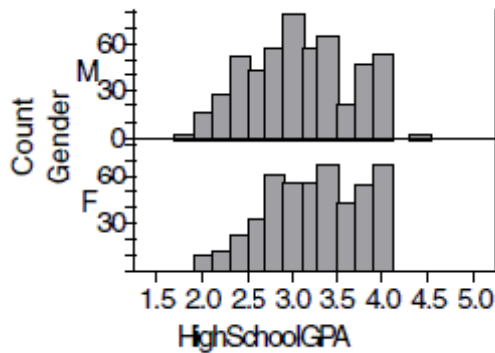
- a. Sketch a histogram for these data.
- b. Find the mean and standard deviation of the plant heights.
- c. Is it appropriate to use the mean and standard deviation to summarize these data? Explain.
- d. Describe the distribution of plant heights.
- 13) All students in a physical education class completed a basketball free-throw shooting event and the highest number of shots made was 32. The next day a student who had just transferred into the school completed the event, making 35 shots. Indicate whether adding the new student's score to the rest of the data made each of these summary statistics increase, decrease, or stay about the same.
- a. mean
- b. median
- c. range
- d. IQR
- e. standard deviation
- 14) Suppose that the student taking 22 credit hours in the data set in the previous question was actually taking 28 credit hours instead of 22 (so we would replace the 22 in the data set with 28). Indicate whether changing the number of credit hours for that student would make each of the following summary statistics increase, decrease, or stay about the same:
- a. mean
- b. median
- c. range
- d. IQR
- e. standard deviation
- 15) On Monday, a class of students took a big test, and the highest score was 92. The next day, a student who had been absent made up the test, scoring 100. Indicate whether adding that student's score to the rest of the data made each of these summary statistics increase, decrease, or stay about the same:
- a. mean
- b. median
- c. range
- d. IQR
- e. standard deviation

13) _____

14) _____

15) _____

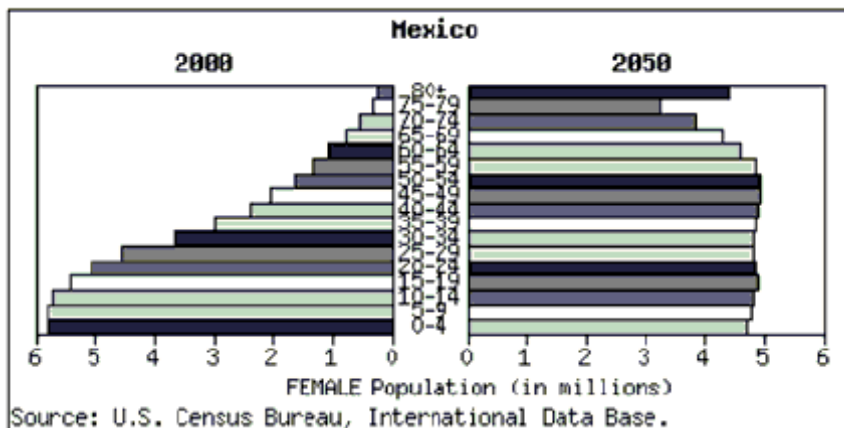
- 16) One thousand students from a local university were sampled to gather information such as gender, high school GPA, college GPA, and total SAT scores. The results were used to create histograms displaying high school grade point averages (GPA's) for both males and females. Compare the grade distribution of males and females. 16) _____



- 17) The five-number summary for midterm scores (number of points; the maximum possible score was 50 points) from an intro stats class is: 17) _____

Min	Q1	Median	Q3	Max
16.5	32	39	43.5	48.5

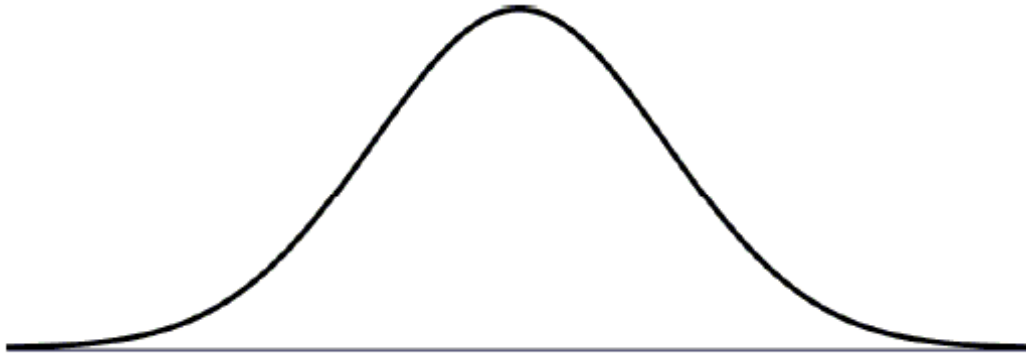
- a. Would you expect the mean midterm score of all students who took the midterm to be higher or lower than the median? Explain.
b. Based on the five-number summary, are any of the midterm scores outliers? Explain.
- 18) At www.census.gov you can create a "population pyramid" for any country. These pyramids are back-to-back histograms. This pyramid shows Mexico's 2000 female population and the census bureau's projection for 2050. Write a few sentences summarizing the changes that are forecast. 18) _____



- 19) The five-number summary for the weights (in pounds) of fish caught in a bass tournament is: 19) _____
- | Min | Q1 | Median | Q3 | Max |
|-----|-----|--------|-----|-----|
| 2.3 | 2.8 | 3.0 | 3.3 | 4.5 |
- a. Would you expect the mean weight of all fish caught to be higher or lower than the median? Explain.
- b. You caught 3 bass weighing 2.3 pounds, 3.9 pounds, and 4.2 pounds. Were any of your fish outliers? Explain.
- 20) The five-number summary for the fuel economy (in miles per gallon) of year 2011 midsize cars is: 20) _____
- | Min | Q1 | Median | Q3 | Max |
|-----|----|--------|----|-----|
| 13 | 21 | 23 | 26 | 50 |
- a. Would you expect the mean gas mileage of all midsize cars to be higher or lower than the median? Explain.
- b. One model of Volkswagon gets 34 mpg, one model of Toyota gets 28 mpg, and one model of Bentley gets 13 mpg. Are any of these cars outliers? Explain.
- 21) Although most of us buy milk by the quart or gallon, farmers measure daily production in pounds. Guernsey cows average 39 pounds of milk a day with a standard deviation of 8 pounds. For Jerseys the mean daily production is 43 pounds with a standard deviation of 5 pounds. When being shown at a state fair a champion Guernsey and a champion Jersey each gave 54 pounds of milk. Which cow's milk production was more remarkable? Explain. 21) _____
- 22) The lifespans of a particular brand of graphing calculator are approximately normally distributed with a mean of 620 days from the purchase date and a standard deviation of 82 days. They will provide a warranty that guarantees a replacement if the calculator stops working within the specified time frame, and are trying to decide what time frame to use. 22) _____
- a. If the company sets the warranty at a year and a half (say 540 days), what proportion of calculators will they have to replace?
- b. The company does not want to have to replace more than 1% of the calculators they sell. What length of time should they set for the warranty?
- c. The company would like to set the warranty for 540 days, and still replace no more than 1% of the calculators sold. Increasing the average life of the calculators is too expensive, but they think they reduce the standard deviation of the lifespans. What standard deviation of lifespans would be needed to make this happen?
- d. Explain what achieving a smaller standard deviation means in this context.

- 23) Owners of a minor league baseball team believe that a Normal model is useful in projecting the number of fans who will attend home games. They use a mean of 8500 fans and a standard deviation of 1500 fans. Draw and clearly label this model.

23) _____



- 24) A statistics teacher gave her class a 15 point quiz. The summary statistics for the students' scores are shown in the table at the right.

24) _____

\bar{x}	10.95 points
s	2.481 points
min	4
Q1	9.5
median	12
Q3	12
max	15

- Notice that the median score and the third quartile are the same. Explain how this can be.
 - One student's parent heaped praise on him for scoring 13, saying it was an amazing score. Comment on whether that praise is deserved using the summary statistics as support.
 - To convert these raw scores to a score out of 100, the teacher multiplies each score by six, then adds 10. (We can debate the wisdom of such a strategy later!). What is the median converted score? And the IQR?
 - What are the mean and standard deviation of the converted test scores?
- 25) Dimes minted in the United States average 2.286 g with a standard deviation of 0.06 g. A couple chemistry students were trying out their teacher's new scale by weighing a bunch of coins. They found a nickel that weighed 5.19 g and a dime that weighed 2.45 g. Which coin was more exceptionally heavy? Explain.

25) _____

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

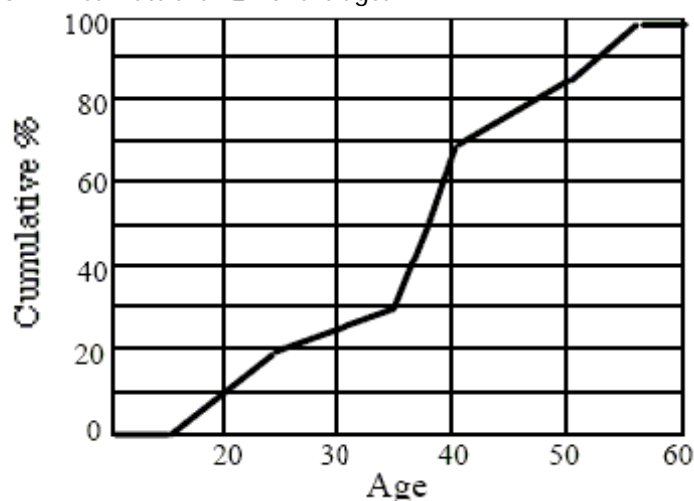
- 26) Suppose that a Normal model described student scores in a history class. Parker has a standardized score (z-score) of +2.5. This means that Parker
- is 2.5 standard deviations above average for the class.
 - none of these
 - has a standard deviation of 2.5.
 - has a score that is 2.5 times the average for the class.
 - is 2.5 points above average for the class.

26) _____

- 27) Which of these variables is most likely to follow a Normal model? 27) _____
- A) eye color
 - B) head circumference
 - C) number of TV sets at home
 - D) hours of homework last week
 - E) number of cigarettes smoked daily

- 28) Environmental researchers have collected rain acidity data for several decades. They want to see if there is any evidence that attempts to reduce industrial pollution have produced a trend toward less acidic rainfall. They should display their data in a(n)... 28) _____
- A) boxplot
 - B) contingency table
 - C) timeplot
 - D) histogram
 - E) bar graph

- 29) The ages of people attending the opening show of a new movie are summarized in the ogive shown. Estimate the IQR of the ages. 29) _____



- A) 5 B) 37 C) 21 D) 30 E) 13
- 30) The advantage of making a stem-and-leaf display instead of a dotplot is that a stem-and-leaf display 30) _____
- A) satisfies the area principle.
 - B) none of these
 - C) shows the shape of the distribution better than a dotplot.
 - D) preserves the individual data values.
 - E) A stem-and-leaf display is for quantitative data, while a dotplot shows categorical data.